Molecular changes in two maize (*Zea mays* L.) synthetics after reciprocal selection with two alternative methods

Ordas B [a], Malvar RA [a], Díaz R [a], Butron A [a]

[a] Misión Biológica de Galicia, CSIC, P.O. box 28, 36080 Pontevedra, Spain.

Communicating author: Ordas B, bordas@mbg.csic.es, telephone: +34986854800, fax: +34986841362

Agronomic evaluations demonstrated that a modification of the classical full-sib reciprocal recurrent selection (RRS-FS) which, in addition to crosses, uses S_2 families evaluation (RRS-FS-S_2) is more efficient than the classical method for developing high yielding crosses between two varieties. The objective of this study was to investigate the changes in genetic diversity and structure after performing RRS-FS and RRS-FS-S_2 selections. RRS-FS-S_2 reduced more the variability, produced more differentiation between cycles of selection derived from the same materials but less between reciprocal populations, and produced a more clear change in the contribution of the parental lines than RRS-FS. On the other hand, the type of selection method did not have a considerable effect on the structure of the populations measured as departure of Hardy-Weinberg (HW) equilibrium at single markers and on linkage disequilibrium (LD) between pairs of markers. We identified some individual markers which were not in HW equilibrium in several populations probably due to genes favouring assortative mating. We also found pairs of markers which increased their LD with selection probably due to epistasis.

Keywords: Hardy-Weinberg equilibrium; linkage disequilibrium; recurrent selection, molecular markers; population genetics; genetic variability

34

35 Introduction

36

37 Interpopulation selection methods are of interest for crops as maize (Zea mays L.)
38 because the commercial varieties are mainly hybrids. Reciprocal recurrent selection
39 (RRS) was proposed to improve the performance of two populations, as well as the
40 cross between them (Comstock et al. 1949). Classical full-sib RRS (RRS-FS) uses
41 interpopulation full-sib progenies (Hallauer and Eberhart 1970). A modification of the
42 RRS-FS procedure was proposed (Moreno-González and Hallauer 1982) which
43 evaluates S2 families in addition to full-sib progenies (RRS-FS-S_2). Two parallel RRS
44 programs (RRS-FS and RRS-FS-S_2) were carried out from the same original
45 populations of maize to assess the relative efficiency of each method (Ordas et al.
46 2012). The evaluation of those RRS programs allowed to conclude that for developing
47 high yielding and stable crosses between two varieties RRS-FS-S_2 is more efficient than
48 RRS-FS.

49 In recurrent selection, there is a tradeoff between the intensity of selection and the
50 effective population size given that the total number of families evaluated is fixed as
51 determined by resources. This, in turn, implies a tradeoff between short term selection,
52 favoured by higher intensities and long term selection favoured by higher effective
53 population sizes which maintain higher levels of variability for longer. The estimation
54 of the change of the genetic variation and structure of populations under selection
55 allows an assessment of the expected progress of selection in the long term. The genetic
56 variability and structure of populations can be efficiently estimated with molecular
57 markers and the information obtained can be used to choose the selection methodology

more appropriate to our breeding objectives. By means of molecular markers, it was found some reduction in genetic variability during the cycles of intrapopulation or interpopulation recurrent selection programs for the number of families usually selected (10-20 selected out of 100-200 evaluated) in those programs; the magnitude of the lost of variability depended of the germplasm and the particular program (Labate et al. 1997; Pinto et al. 2003; Butron et al. 2005; Romay et al. 2012; Peña-Asin et al. 2013). In addition to a lost of variability, during RRS-FS selection, a change on the structure of the populations, affecting the linkage disequilibrium (LD) and the genetic distance between reciprocal population, was observed (Labate et al. 1997; Romay et al. 2012; Peña-Asin et al. 2013). However, there is not information about the effect of RRS-FS-S_2 on the molecular genetic variation and structure of the selected populations. Therefore, the objective of this study is to investigate the changes in molecular variation and population structure (Hardy-Weinberg (HW), LD and relative contribution of the parental lines to the synthetics) following RR-FS and RR-FS-S2 selection, comparing the changes resulting from both methods

Material and methods

Plant material and simple sequence repeat (SSR) genotyping

The populations analyzed in this study were derived from two parallel programs of RRS starting from EPS20 (a Reid synthetic) and EPS21 (a no Reid synthetic) (Supplementary Table 1). Two cycles of RRS-FS and RRS-FS-S_2 (Supplementary Fig 1) were carried out as described in Ordas et al. (2012), resulting in 10 populations (the 2 original populations and 8 derived populations-2 synthetics × 2 cycles × 2 methods-). Forty-eight individuals were sampled from each population for DNA extraction. The DNA was extracted following the methodology of Liu and Whittier (1994). Fifty two SSRs distributed throughout the genome were examined. The SSRs were selected because in previous studies performed adequately in multiple populations (Butron et al. 2005; Romay et al. 2012). The SSR products were separated by capillary electrophoresis using 1x Tris base, boric acid (5.5 g l^{-1}) and ethylendialminetetraacetic acid (EDTA) (2 mM) on a polymerase TAQ (30,000 ud). A Beckman Coulter CEQ 8800 Genetic analysis system (Beckman Coulter Inc.) was used for fragment separation and identification.

Statistical analysis

The following genetic variability parameters were calculated for all populations: percentage of polymorphic loci, number of alleles per locus, observed heterozygosity and fixation index. To study the change in relationships between populations during selection, genetic distance between populations was calculated using Nei's distance (Nei

1978). Hardy-Weinberg equilibrium at individual markers and the linkage disequilibrium for pair of markers were tested with an exact test based on the multinomial distribution which determines the probabilities of all possible samples of the same size as the observed sample assuming the hypothesis is true (Weir 1996). All analyses were performed with Genetic Data Analysis (Lewis and Zaykin 2001).

Results

Genetic diversity and genetic distances between populations

For EPS20 and EPS21, the variability measured as percentage of polymorphic loci, number of alleles per locus and observed heterozygosity was generally higher in the original population than in selected cycles of RRS-FS and RRS-FS-S_2 (Table 1). The values were lower in the cycles of RRS-FS-S_2 than in the cycles of RRS-FS for most of the variability parameters estimated, particularly for the cycles derived from EPS21. The genetic distance between reciprocal populations increased with both types of RRS (Table 2), being slightly higher for RRS-FS than for RRS-FS-S_2. The genetic distance between the original populations, EPS20 and EPS21, and their improved cycles of selection increased with RRS-FS and RRS-FS-S_2. The increment was higher for materials derived from EPS21 than for EPS20 and for RRS-FS-S_2 than for RRS-FS.

The distance between the synthetic parental lines and the original composites varied from 0.22 to 0.48 between lines in EPS20, but it was close to 0.5 for most lines in EPS21 (Table 3). RRS-FS maintained the genetic contribution of most parental lines observed in the original cycles, while RRS-FS-S_2 changed the genetic contributions of the lines in both populations.

Hardy-Weinberg (HW) and linkage disequilibrium (LD)

The fixation index (f), that is, the excess of homozygosity, averaged over all loci and populations was 0.02. The sign of the fixation index, averaged over loci, was positive for eight of the populations, while its magnitude varied between -0.04 and 0.08 between populations (Table 1). The number of loci not in HW equilibrium varied between 10 and 15 (about 20-30 % of the markers) in most populations. The cycle 2 of selection tended to have few loci in HW disequilibrium and less proportion of polymorphic loci. The number of markers in disequilibrium was similar for the two types of RRS (data not shown). Sixteen markers, located in all chromosomes except chromosome 8, were not in HW equilibrium in three or more of the populations (Table 4). Eleven of those markers had an excess of homozygosity in most populations, while fourteen of them had a positive fixation index, averaged over populations. The markers bnlg1347 (bin 1.10), phi114 (7.03), and umc1453 (10.04) had the highest distortion toward homozygosity excess (fixation index close to 0.30) and the marker umc1505 (9.05) toward heterozygosity excess (fixation index close to -0.30).

141 40 % and 21 % pairs of markers were in linkage disequilibrium (LD), tested at 5 % of
142 significance, in EPS20 and EPS21, respectively. The number of pairs in LD decreased
143 with both selections (RRS-FS and RRS-FS $-S_2$) in EPS20, but increased in EPS21
144 (Table 5).. For some chromosomes the number of pairs of markers with LD changed
145 without any appreciable tendency across the cycles of selection (Table 5). However, for
146 chromosome 9, the number of pairs of markers in LD was consistently low in most
147 populations. LD in chromosome 1 was high in EPS20 and populations derived from it
148 (about 50% of the pairs in LD), but low in the EPS21 and related populations (about 10-
149 20%). LD at chromosome 4 was also high (60%) in EPS20, and was reduced with
150 selection, while LD was low in EPS21 (0%) and increased in the cycles of selection
151 derived from it. LD at chromosome 2 was high (50%) across EPS21 cycles, but low
152 (20%) in EPS20 (although LD was higher in the cycles of selection, particularly C2). At
153 chromosome level, the number of pairs in LD in the selection cycles of RRS-FS was
154 similar to the number of pairs in RRS-FS-S_2.For example, both RRS-FS and RRS-FS-S_2
155 maintained a high level of LD at chromosome 1 in EPS20 and a low level in EPS21. At
156 local level, we observed that some markers were involved in higher or lower number of
157 pairs with significant LD than other markers, for example, phi056 (1.01) in EPS20-
158 RRS-FS-S_2 and umc1466 (4.08) in EPS21-RRS-FS-S_2 (Table 6). Also, about 25 %, 90
159 % and 100 % of pairs in which phi109275 (1.03) and phi114 (7.03) were involved had a
160 significant LD in EPS20C0, EPS20FSC1 and EPS20FSC2 (data not shown), but the
161 total percentage of pairs in LD, considering all markers, in these populations were 40 %,
162 37 % and 20 %. For these markers and also for markers umc1165 (2.01), umc1453
163 (10.04) and umc1930 (10.04), the number of pairs in LD tended to be higher in the
164 selected cycles than in the original populations (data not shown). The marker umc1466
165 had in EPS20 similar change in allele frequencies in both selection programs, but

166 different change in LD, while umc1963 had in EPS21 similar change in LD in both
167 programs, but different change in allele frequencies (Table 6).

168

169 Discussion

170

171 Genetic diversity and genetic distances between populations

172

173 The variability of the original populations was in the range of magnitude of original
174 population of other RRS programs (Labate et al. 1997; Pinto et al. 2003; Romay et al.
175 2012; Peña-Asin et al. 2013). According to different parameters (percentage of
176 polymorphic loci, number of alleles per locus and observed heterozygosity), EPS21 has
177 higher variability than EPS20. This was expected as several parental inbreds of EPS20
178 are related and derived from B14 or WF9, while the origin of the parental inbreds of
179 EPS21 is more diverse: open pollinated varieties from Northwestern Spain, Italy France,
180 etc (Butrón et al. 2003, 2009). Ordas et al. (2012) already found that EPS21 had higher
181 additive variance than EPS20 in the evaluation of several S_2 families derived from both
182 populations. The decrease in variability per cycle of selection found in our selection
183 program was in similar range (0-2%, 2-8%, and 2-6%, for percentage of polymorphic
184 loci, number of alleles per locus and observed heterozygosity, respectively) to other
185 intra and inter population selection programs (Butron et al. 2005; Romay et al. 2012;
186 Peña-Asin et al. 2013). This decrease could be due to the effect of selection increasing
187 and decreasing the frequency of favourable and unfavourable alleles, respectively, and
188 the frequency of neutral alleles linked to them. The frequency changes could be also due

189 to the effect of genetic drift which is originated for the reduced number of families
190 selected each generation. Otherwise, the reduced number of families is mandatory if a
191 high intensity of selection has to be maintained with an approachable number of
192 families under evaluation. Regarding the comparison between the two methods of
193 selection, the decrease in variability was more evident in RRS-FS-S_2 which is also more
194 efficient for improving grain yield, according to theoretical predictions and empirical
195 data (Moreno-González and Hallauer 1982; Ordas et al. 2012). The higher efficiency of
196 RRS-FS-S_2 compared to RRS-FS implies a higher fixing rate of favourable alleles and
197 neutral alleles linked to them which could account for a reduction in variability. The
198 RRS-FS-S_2 was effective by increasing the yield of EPS20 and EPS20 × EPS21, but not
199 the yield of EPS21 although this population presented a high variability (Ordas et al.
200 2012). In RRS-FS-S_2, EPS21 had higher loss of variability than EPS20, according to our
201 molecular data, which could be indicative of higher genetic drift and inbreeding
202 depression. The inbreeding depression can counterbalance the favourable effect of
203 selection. Peña-Asin et al. (2013) also found that one of the two reciprocal populations
204 did not response to RRS; that population had higher loss of variability (according to
205 molecular markers) and higher inbreeding depression (estimated from crosses to the
206 original population).

207

208 In agreement with the increase of heterosis found in the phenotypic evaluation of the
209 RRS program (Ordas et al. 2012), the genetic distance between the original populations
210 increased with RRS. This result was also found in other RRS (Hinze et al. 2005; Romay
211 et al. 2012). The increment of genetic distance between reciprocal populations was
212 slightly higher in RRS-FS than in RRS-FS-S_2 which also is in concordance with a
213 higher increment in heterosis found in RRS-FS (Ordas et al. 2012). The criteria of

214 selection in RRS-FS was exclusively based on the performance of the crosses between
215 families from the two reciprocal populations which could favour more the heterosis than
216 the criteria of selection in RRS-FS-S_2 which also included information of the S_2 families
217 (Moreno-González and Hallauer 1982). On the other hand, the distance between the
218 original and the improved populations was higher with RRS-FS-S_2 probably due to a
219 more efficient increment of favourable alleles with additive effects.

220

221 The genetic distances of the synthetic parental lines to the original and selected
222 populations were estimated to quantify the contribution of each line to the populations
223 and to measure the effect of both types of selection on the relative contribution of the
224 lines. EPS20 was developed with 5 lines derived from B14 and 3 lines derived from
225 WF9. The distance between the lines and the original synthetic reflects this structure.
226 The lines derived from B14, particularly those with higher proportion of this line, were
227 more close to the synthetic (distance=0.2-0.3) than the lines derived from WF9
228 (distance=0.4-0.5) (Butron et al. 2003; 2009). On the contrary, EPS21 was developed
229 from lines from different origin and the distance between them and the original
230 synthetic was similar for most of the lines (distance=0.5, approximately). In agreement
231 with Labate et al. (1997), RRS-FS maintained the contribution observed in the original
232 cycles for most of the lines, except EP17 in EPS21. The loss of contribution of this line
233 was also observed in RRS-FS-S_2, but at a lower level. It is possible that EP17 had lower
234 amount of favourable alleles or complementary alleles to EPS20 than other lines of
235 EPS21. Contrary to RRS-FS, RRS-FS-S_2 changed the genetic contributions of the lines
236 in both populations. In EPS20, RRS-FS-S_2 increased the distance between the synthetic
237 and several of the B14 derived lines making that most lines had similar distances with
238 the synthetic (about 0.4-0.5). Therefore, selection could had favoured the WF9 derived

239 lines at the expense of the B14 derived lines, probably due to the best adaptation of
240 WF9 derived lines to Atlantic conditions. In EPS21, RRS-FS-S_2 increased the distance
241 between the lines and the synthetic, particularly with the lines EP53, PB130 and F473.
242 Given that the increments that showed the population cross with RRS-FS-S_2 were
243 similar to the increments shown by EPS20 (Ordas et al. 2012) the authors suggested that
244 RRS-FS-S_2, at difference of RRS-FS, was able to manipulate the additive effects,
245 particularly those present in EPS20. The molecular data were consistent with this
246 hypothesis, and suggest, besides, that those favourable additive effects could come from
247 the WF9 lines.

248

249 Hardy-Weinberg and linkage disequilibrium

250

251 The percentage of loci not in Hardy-Weinberg equilibrium in our populations was
252 similar to those found in other selection experiments (Kahler et al. 1986; Labate et al.
253 2000). In agreement with several selection experiments (Kahler et al. 1986; Labate et al.
254 2000; Butron et al. 2005; Hinze et al. 2005; Romay et al. 2012) we observed an excess
255 of homozygosity. Some hypotheses have been proposed to explain the excess of
256 homozygous, including reduced sampling size, positive assortative mating and
257 genotyping errors. However, the magnitude of the fixation index in our selection
258 program seems to be too small to have important implications. In general, we conclude
259 that the sampling size or the assortative mating did not greatly affect the structure of the
260 populations, at least, when the equilibrium is estimated at individual loci. Although the
261 excess of homozygous was not important when the total effect was measured as the
262 average value of all markers, it could be relevant in specific regions of the genome. This

263 is especially evident for three markers (bnlg1347, phi114, and umc1453) which had a
264 positive fixation index in 9 of the populations, a significant departure from Hardy-
265 Weinberg in 5 or 6 populations, and a large value of the parameter (f=0.3). We expected
266 that distortions due to sampling size would not be restricted to local regions, but spread
267 along the genome. In distortions due to sampling size, we did not expect either a high
268 consistency across populations as we found in some of the markers. The most likely
269 cause of consistent distortion in some of the markers is the positive assortative mating
270 due to differences in flowering time between the families. The marker bnlg1347 had a
271 significant departure of Hardy-Weinberg equilibrium toward homozygosity not only in
272 5 of our populations, but also in one of the populations analyzed by Butron et al. (2005)
273 and Romay et al. (2012). However, the most consistent departure from Hardy-Weinberg
274 equilibrium occurred in umc1453 (10.04): in 6 out of 10 populations in our experiment
275 (f=0.3), in 3 out of 3 in Butron et al. (2005) (f=0.3), and 4 out of 6 in Romay et al.
276 (2012) (f=0.2). This marker is located in chromosome 10 at the bin position 10.04
277 where one of main QTLs for flowering time has been found; the importance of this QTL
278 is due not only to its large effect, but also due to its consistency in different genetic
279 backgrounds (Ducrocq et al. 2009 and references therein). The coincidence of a large
280 QTL for flowering time and a significant excess of homozygous in the same region
281 supports the hypothesis of genetic distortion due to positive assortative mating (Butron
282 et al. 2005).

283 The high level of LD in the original synthetics was probably generated during the
284 development of the synthetics due to the cross of inbred lines of different origin. The
285 effect of RRS selection on the number of pairs of markers in LD was different in EPS20
286 (about 10 % less LD in C2) and EPS21 (about 5 % more LD in C2). Similar results
287 were obtained in other RRS programs in which one population increased and the other

decreased the pairs of markers in LD (Labate et al. 2000; Romay et al. 2012; Peña-Asin et al. 2013). The both types of reciprocal selections have similar effects on the total number of pairs in LD and on the number of markers not in HW equilibrium, suggesting that the type of recurrent selection seems not to have a large effect in the structure of the populations. We found some heterogeneity between chromosomes in the number of pairs of markers in LD, for example, the chromosome 9 had a lower number of pairs in LD than other chromosomes across different populations. This was probably due to variation in chromosome specific recombination rates which was reported in maize by Bauer et al. (2013). We expected in regions with higher recombination rates lower LD. Bauer et al. (2013) found that the chromosome 9 had the highest recombination rate which is in concordance with our results. We found some differences in LD between EPS20 (a dent synthetic) and EPS21 (a flint synthetic), particularly for chromosomes 1, 2 and 4, which is also in concordance with results of Bauer et al. (2013) who found heterogeneity in the recombination profiles of Dent and Flint populations for some chromosomes, including chromosomes 2 and 4. We found some heterogeneity in LD not only at chromosome level but also at local level as some markers were involved in higher or lower number of pairs with significant LD than other markers. Thus, although the total percentage of pairs (counting all markers) with LD decreased with selection in EPS20, some particular markers had higher number of LD in the selected cycles of EPS20 than in the original population. The most evident increment in LD was observed for phi114 and phi109275 in RRS-FS which changed from a lower number of pairs in LD than the average (20-30% vs. 40%) in EPS20 to a higher number than the average (90-100% vs. 20-37%) in the selected cycles. Epistasis acting at local level could have generated this increment in LD in some markers of EPS20 and also could play a role in the increment of LD found in EPS21 across the genome. Epistasis has been proposed as

313 a possible explanation for the increasing of LD found in other RRS programs (Romay et
314 al. 2012; Peña-Asin et al. 2013). LD can be also generated by random drift in small
315 populations, but only between loosely linked loci (Labate et al. 2000) which is not the
316 case in our experiment. In some markers, a similar change in allele frequency did not
317 imply a similar change in LD, and vice versa, a similar change in LD did not imply a
318 similar change in allele frequency which could be indicative that the change in allele
319 frequency did not greatly affect the LD.

320

321 The RRS-FS-S_2 method produces higher improvement of the population cross and of
322 one of the populations per se than the RRS-FS method. The molecular data reported
323 here gives complementary information about the characteristics of each method. In our
324 experiment, RRS-FS-S_2 reduced more the variability, produced more differentiation
325 among populations derived from the same materials but less differentiation between
326 reciprocal populations, and produced a more clear change in the contribution of the
327 parental lines than RRS-FS. These data are consistent with the hypothesis that RRS-FS-
328 S_2 is more efficient than RRS-FS because is able to manipulate the additive effects
329 (Ordas et al. 2012). On the other hand, the type of selection method did not have a
330 considerable effect in the structure of the populations measured as departure of HW
331 equilibrium at single markers or LD between pairs of markers. The generalization of our
332 conclusions has to be made with caution as our results are based on one selection
333 process without replication. We identified some individual markers which were not in
334 HW equilibrium in several of the populations probably due to genes favouring
335 assortative mating. We also found pairs of markers which increased their LD with
336 selection probably due to epistasis.

337 Acknowledgments

338 Research was supported by the Spanish Plan of Research and Development (AGL2012-
339 33415). Bernardo Ordas acknowledges a Ramon y Cajal contract from the Ministry of
340 Economy and Competitiveness

341

342 Bauer E, Falque M, Walter H et al. (2013) Intraspecific variation of recombination rate
343 in maize. Gen Biol 14:R103

344 Butron A, Tarrío R, Revilla P, Ordás A, Malvar RA (2005) Molecular changes in the
345 maize composite EPS12 during selection for resistance to pink stem borer. Theor
346 Appl Genet 110:044-1051

347 Butron A, Revilla P, Romay MC, Ordas A, Malvar RA (2009) Causes of agronomic
348 differences between synthetics developed by the random and convergent cross
349 methods. Fiel Crop Res 110:229-234

350 Comstock RE, Robinson HF, Harvey PH (1949) A breeding procedure designed to
351 make maximum use of both general and specific combining ability. Agron J
352 41:360-367

353 Ducrocq S, Giauffret C, Madur D, Combes Valerie, Dumas F, Jouanne S, Coubriche D,
354 Jamin P, Moreau L, Charcosset A (2009) Fine mapping and haplotype structure
355 analysis of a major flowering time quantitative trait locus on maize chromosome
356 10. Genetics 183:1555-1563

357 Hallauer AR, Eberhart SA (1970) Reciprocal full-sib selection. Crop Sci 10:315-316

358 Hinze LL, Kresovich S, Nason JD, Lamkey KR (2005) Population genetic diversity in a
359 maize reciprocal recurrent selection program. Crop Sci 45:2435-2442

360 Kahler AI, Hallauer AR, Gardner CO (1986) Allozyme polymorphisms within and
361 among open-pollinated and adapted exotic populations of maize. Theor Appl Genet
362 72:592-601

363 Labate JA, Lamkey KR, Lee M, Woodman WL (1997) Molecular genetic diversity after
364 reciprocal recurrent selection in BSSS and BSCB1 maize populations. Crop Sci
365 37:416-423

366 Labate JA, Lamkey KR, Lee M, Woodman WL (2000) Hardy-Weinberg and linkage
367 equilibrium estimates in the BSSS and BSCB1 random mated populations. Maydica
368 45:243-255

369 Lewis PO, Zaykin D (2001) Genetic Data Analysis: Computer program for the analysis
370 of allelic data. Versión 1.0 (d16c).
371 http://lewis.eeb.uconn.edu/lewishome/software.html

372 Liu YG, Whittier RF (1994) Preparation of megabase plant DNA from nuclei in agarose
373 plugs and microbeads. Nucleic Acids Research 22:2168-2169

374 Moreno-Gonzalez J, Hallauer AR (1982) Combined S_2 and crossbred family selection in
375 full-sib reciprocal recurrent selection. Theor Appl Genet 61:353-358

376 Nei M (1978) Estimation of average heterozigosity and genetic distance from a small
377 number of individuals. Genetics 89:583-590

378 Ordas B, Butron A, Alvarez A, Revilla P, Malvar RA (2012) Comparison of two
379 methods of reciprocal selection in maize (*Zea mays* L.). Theor Appl Genet
380 124:1183-1191

381 Peña-Asin J, Alvarez A, Ordas B (2013) Molecular changes during intra and inter
382 recurrent selection of two populations of maize: one adapted and one non adapted
383 to the selection environment. Euphytica 193:359-367

Pinto LR, Vieira MCL, de Souza CL, de Souza AP (2003) Genetic diversity assessed by microsatellites in tropical maize populations submitted to a high-intensity reciprocal recurrent selection. Euphytica 134:277-286

Romay MC, Butron A, Ordas A, Revilla P, Ordas B (2012) Effect of recurrent selection on the genetic structure of two broad-based Spanish maize populations. Crop Sci 52:1493-1502

Weir BS (1996) Genetic data analysis II. Sinauer Associates, Sunderland, MA, USA

Table 1 Percentage of polymorphic loci (Po), number of alleles per locus (A), observed heterozygosity (Ho) and fixation index (f) for the maize synthetics EPS20 and EPS21 and the cycles selected from them by RRS-FS and RRS-FS-S_2

Populations	Po	A	Ho	f
EPS20	0.96	3.06	0.44	0.06
EPS20(RRS-FS)C1	0.96	3.51	0.49	0.00
EPS20(RRS-FS)C2	0.91	2.92	0.39	0.04
EPS20(RRS-FS-S_2)C1	0.92	3.00	0.43	0.05
EPS20(RRS-FS-S_2)C2	0.92	2.75	0.43	-0.04
EPS21	1.00	3.81	0.56	0.04
EPS21(RRS-FS)C1	1.00	3.66	0.54	0.01
EPS21(RRS-FS)C2	1.00	3.68	0.55	0.00
EPS21(RRS-FS-S_2)C1	1.00	3.58	0.50	0.08
EPS21(RRS-FS-S_2)C2	0.98	3.38	0.44	0.00

399 **Table 2** Genetic distances betweenthe maize synthetics EPS20 and EPS21 and the cycles selected from them by RRS-FS and RRS-FS-S_2

Populations	EPS20	EPS20(RRS-FS)C1	EPS20(RRS-FS)C2	EPS20(RRS-FS-S_2)C1	EPS20(RRS-FS-S_2)C2	EPS21	EPS21(RRS-FS)C1	EPS21(RRS-FS)C2	EPS21(RRS-FS-S_2)C1
EPS20(RRS-FS)C1	0.04								
EPS20(RRS-FS)C2	0.07	0.04							
EPS20(RRS-FS-S_2)C1	0.04	0.04	0.06						
EPS20(RRS-FS-S_2)C2	0.09	0.07	0.10	0.05					
EPS21	0.30	0.32	0.39	0.32	0.35				
EPS21(RRS-FS)C1	0.36	0.36	0.44	0.37	0.40	0.05			
EPS21(RRS-FS)C2	0.33	0.33	0.41	0.35	0.38	0.07	0.09		
EPS21(RRS-FS-S_2)C1	0.30	0.32	0.38	0.32	0.37	0.04	0.08	0.06	
EPS21(RRS-FS-S_2)C2	0.34	0.34	0.43	0.35	0.37	0.16	0.11	0.20	0.22

400

401

402 **Table 3** Genetic distances between the parental lines and the original and improved synthetics

	CM109	**CM139**	**CM151**	**A634**	**A639**	**A652**	**A664**	**W64A**
EPS20	0.22	0.22	0.48	0.29	0.41	0.45	0.35	0.44
EPS20(RRS-FS)C1	0.22	0.22	0.48	0.28	0.38	0.46	0.41	0.42
EPS20(RRS-FS)C2	0.24	0.24	0.46	0.31	0.35	0.52	0.44	0.47
EPS20(RRS-FS-S_2)C1	0.27	0.27	0.49	0.28	0.40	0.42	0.33	0.46
EPS20(RRS-FS-S_2)C2	0.36	0.36	0.47	0.31	0.45	0.46	0.44	0.47
	EP17	**EP43**	**EP53**	**PB60**	**PB130**	**F473**	**CO125**	**A509**
EPS21	0.51	0.46	0.50	0.33	0.50	0.50	0.45	0.44
EPS21(RRS-FS)C1	0.56	0.46	0.65	0.37	0.59	0.51	0.41	0.53
EPS21(RRS-FS)C2	0.70	0.55	0.52	0.41	0.50	0.41	0.51	0.47
EPS21(RRS-FS-S_2)C1	0.67	0.49	0.48	0.43	0.50	0.50	0.43	0.46
EPS21(RRS-FS-S_2)C2	0.61	0.56	0.74	0.36	0.75	0.70	0.58	0.59

403

404 **Table 4** For each marker, the number of populations in which the marker is not in HW
405 equilibrium and has positive and negative fixation indexes. Only those markers with
406 significant departure from HW equilibrium in more than two populations are shown.
407 For each marker the average fixation index over populations is also shown

Marker	Location (bin)	Significant HW (population number)	Positive fixation index (population number)	Negative fixation index (population number)	Average fixation index
Umc1222	1.02	5	7	3	0.10
Phi109275	1.03	3	5	5	0.05
Bnlg1347	1.10	5	9	1	0.31
Umc1165	2.01	5	8	2	0.16
Umc1185	2.03	4	5	5	0.02
Phi036	3.04	4	6	4	0.02
Umc1466	4.08	3	5	5	0.09
Umc1822	5.05	4	8	2	0.06
Phi128	5.07	4	8	2	0.19
Bnlg1154	6.05	3	8	2	0.16
Bnlg1740	6.07	5	6	4	0.11
Umc1653	6.07	3	5	0	0.19
Phi114	7.03	5	9	0	0.27
Umc1505	9.08	5	1	9	-0.28
Umc1453	10.04	6	9	1	0.32
Umc1930	10.04	5	4	6	-0.07

Table 5 Percentage of pairs of markers in linkage disequilibrium within each chromosome

Chromosome	EPS20	EPS20(RRS-FS)C1	EPS20(RRS-FS)C2	EPS20(RRS-FS-S_2)C1	EPS20(RRS-FS-S_2)C2	EPS21	EPS21(RRS-FS)C1	EPS21(RRS-FS)C2	EPS21(RRS-FS-S_2)C1	EPS21(RRS-FS-S_2)C2
1	0.50	0.46	0.50	0.46	0.57	0.11	0.11	0.21	0.07	0.11
2	0.18	0.50	0.11	0.43	0.29	0.54	0.39	0.57	0.54	0.43
4	0.58	0.14	0.06	0.06	0.36	0.00	0.19	0.19	0.03	0.14
5	0.50	0.80	0.20	0.50	0.10	0.30	0.10	0.30	0.3	0.00
6	0.48	0.48	0.29	0.29	0.43	0.38	0.57	0.52	0.71	0.81
7	0.00	0.33	0.67	0.33	0.00	0.33	0.33	0.33	0.67	0.00
8	0.83	0.33	0.00	0.17	0.17	1.00	0.17	0.00	0.17	0.33
9	0.00	0.40	0.20	0.90	0.10	0.10	0.00	0.00	0.20	0.00
10	0.33	0.67	0.00	0.67	0.67	0.33	0.00	0.67	0.67	1.00
Average within [a]	0.42	0.41	0.21	0.35	0.35	0.26	0.25	0.32	0.30	0.29
Average between [b]	0.39	0.36	0.19	0.34	0.31	0.23	0.20	0.28	0.30	0.25

[a] Average percentage of pairs of markers within the same chromosome in linkage disequilibrium

[b] Average percentage of pairs of markers from different chromosomes in linkage disequilibrium

411 **Table 6** Percentage of pairs of markers in linkage disequilibrium (LD) in the original synthetics and final cycles of selection for markers with

412 high (>98 %) or low (<2 %) percentage of pairs in LD in the original synthetics. For those markers the allele frequencies are also shown

	EPS20	EPS20(RRS-FS)C2	EPS20(RRS-FS-S_2)C2	EPS21	EPS21(RRS-FS)C2	EPS21(RRS-FS-S_2)C2
phi056 (1.01)						
Marker pairs in LD (%)	23	6	35	2	21	12
A allele	0.00	0.00	0.00	0.12	0.27	0.00
B allele	0.00	0.00	0.00	0.16	0.21	0.00
C allele	0.65	0.17	0.09	0.46	0.40	0.07
D allele	0.25	0.56	0.88	0.24	0.10	0.93
E allele	0.08	0.27	0.02	0.00	0.00	0.00
umc1466 (4.08)						
Marker pairs in LD (%)	100	13	98	23	21	6
A allele	0.31	0.00	0.05	0.93	0.96	0.71
B allele	0.69	1.00	0.95	0.07	0.04	0.29
umc1963 (4.04)						
Marker pairs in LD (%)	40	12	17	2	27	23
A allele	0.87	0.98	1.00	0.91	0.71	1.00
B allele	0.00	0.00	0.00	0.02	0.08	0.00
C allele	0.11	0.02	0.00	0.04	0.017	0.00
D allele	0.02	0.00	0.00	0.02	0.04	0.00

413